

Improving Anatomical Continuity in Synthetic Left Atrial LGE-MRI Using Inter-Slice Attention

Miao Zhao^{1*}[0009-0003-5555-7965], Shibo Li¹[0000-0002-2369-336X], Xu Chen¹[0000-0001-9925-2598], Sayed Al-Aidarous¹[0000-0002-7044-9423], Shohreh Honarbakhsh¹[0000-0003-0081-5739], Venet Osmani¹[0000-0001-7306-2972], Gregory Slabaugh¹[0000-0003-4060-5226], and Caroline Roney¹[0000-0001-6809-0928]

¹Queen Mary University of London, London, United Kingdom
miao.zhao@qmul.ac.uk

Abstract. Atrial fibrillation (AF) research increasingly relies on virtual clinical trials and digital twins, yet the development of large populations of atrial models with anatomical variability is constrained by the limited availability of late gadolinium-enhanced magnetic resonance imaging (LGE-MRI). Recent mask-conditioned diffusion models have, therefore, emerged as a practical approach for generating realistic LGE-MRI appearance from available anatomical labels, thereby producing paired image-label data for image-based cardiac analysis. However, for digital twin applications, anatomical realism alone is insufficient, as accurate three-dimensional reconstruction also requires consistent structural transitions across adjacent slices within a volumetric image stack. Most existing diffusion models generate slices independently and consequently fail to preserve anatomical continuity along the z-axis, which may compromise volumetric reconstruction quality. We propose Inter-Slice Adaptively Distilled ControlNet (IS-ADC), a continuity-aware diffusion framework that incorporates neighbouring slice information through a teacher-only Inter-Slice Attention module. The learned continuity priors are subsequently transferred to a mask-conditioned student branch via adaptive distillation, improving volumetric consistency while requiring only a modest increase in training cost and no additional inference-time computation. Experiments on the Left Atrial Segmentation Challenge 2018 dataset demonstrate that IS-ADC achieves better image fidelity and downstream segmentation performance than both the original ADC framework and a three-slice-input baseline. Furthermore, IS-ADC improves inter-slice anatomical continuity and structural integrity, producing more coherent three-dimensional left atrial reconstructions.

Keywords: Atrial fibrillation · Diffusion models · Inter-slice consistency.

1 Introduction

Atrial fibrillation (AF) is the most prevalent sustained cardiac arrhythmia, with considerable differences between patients in both disease characteristics and

treatment response [1, 2]. Consequently, identifying the optimal therapy for individual patients remains a significant clinical challenge. Virtual clinical trials (VCTs) and cardiac digital twins have emerged as promising tools for evaluating treatment strategies *in silico*, reducing the cost and burden compared with conventional clinical studies [3]. However, the reliability of these frameworks depends critically on anatomically accurate and structurally consistent cardiac models, as even minor geometric inaccuracies may propagate into downstream electrophysiological analyses and simulation outcomes [4].

Patient-specific cardiac digital twins require large LGE-MRI cohorts, yet image acquisition and annotation remain costly and are further constrained by privacy regulations. These images are commonly used to reconstruct patient-specific cardiac models for electrophysiological simulation and virtual treatment planning [14, 15].

Although anatomical masks provide explicit geometric information, they lack the realistic LGE-MRI appearance required for many downstream image analysis tasks, including intensity distributions, surrounding anatomical context, and acquisition-related variability. Consequently, mask-conditioned image synthesis has emerged as an effective approach for generating realistic image-label pairs when annotated LGE-MRI data are limited.

To address this limitation, mask-conditioned diffusion models, particularly ControlNet [5] and ADC [6], have emerged as effective approaches for generating realistic LGE-MRI from known anatomical masks while preserving patient-specific geometry [7, 8].

Despite these advances, most diffusion-based medical image synthesis methods operate in a slice-wise manner and fail to model inter-slice anatomical continuity. This limitation becomes particularly important when slices are assembled into three-dimensional cardiac models. Although three-dimensional diffusion models [8, 11] can capture the volumetric context, their substantial computational and memory requirements often limit practical adoption, particularly for high-resolution cardiac imaging. As a result, incorporating information from neighbouring slices has emerged as a practical strategy for improving volumetric consistency while maintaining computational efficiency.

To mitigate this limitation, recent studies have explored the use of neighbouring-slice context. Kumar et al. [9] introduced a cross-slice attention mechanism for 2.5D medical image segmentation, demonstrating the effectiveness of attention-based feature interaction for capturing volumetric context. Inter-slice residual learning has similarly been used to aggregate contextual information across adjacent slices in volumetric segmentation [10]. Similar ideas have also been explored in diffusion-based medical image synthesis, where neighbouring slices are exploited to improve volumetric consistency [12]. Related concepts have further emerged in video diffusion models, where sparse temporal attention enables information exchange across neighbouring frames and improves temporal consistency [13].

Nevertheless, despite these encouraging advances, diffusion-based cardiac image synthesis still lacks an effective mechanism for modelling feature-level depen-

dependencies between neighbouring anatomical slices. To address this limitation, we propose an Inter-slice Adaptively Distilled ControlNet (IS-ADC), which incorporates an Inter-Slice Attention (ISA) module into the baseline ADC framework. Inspired by the sparse attention mechanism [13] originally introduced for video generation, ISA adapts temporal attention to model inter-slice anatomical dependencies in LGE-MRI volumes. Unlike video generation, where attention captures motion consistency across consecutive frames, ISA enables the target slice to leverage anatomical information from neighbouring LGE-MRI slices, allowing image generation to be guided by their anatomical shapes and structural variations. This design preserves cross-slice anatomical continuity while remaining fully compatible with the original 2D latent diffusion architecture and avoiding the computational overhead of full 3D diffusion models.

Our contributions are summarised as follows: (1) We propose IS-ADC, an extension of Adaptively Distilled ControlNet that leverages contextual information from neighbouring slices through an Inter-Slice Attention (ISA) module, improving volumetric anatomical consistency without additional inference cost. (2) We establish a comprehensive evaluation framework incorporating image fidelity, segmentation, inter-slice continuity and 3D reconstruction metrics. (3) Extensive experiments demonstrate that IS-ADC generates anatomically realistic LGE-MRI volumes with improved inter-slice consistency, supporting more reliable three-dimensional reconstruction for cardiac digital twin applications.

2 Methods

Fig. 1 illustrates the proposed Inter-Slice Adaptively Distilled ControlNet (IS-ADC) framework. Built upon Adaptively Distilled ControlNet (ADC) [6], IS-ADC introduces a teacher-only Inter-Slice Attention (ISA) module to model anatomical continuity across neighbouring LGE-MRI slices. During training, three consecutive slices are jointly processed, enabling the target slice to aggregate contextual information from adjacent anatomy through inter-slice attention. The enhanced features are incorporated into the teacher branch and subsequently transferred to the student branch through adaptive distillation. During inference, only the student branch is retained, preserving the computational efficiency of ADC while improving volumetric anatomical consistency of synthesised LGE-MRI.

2.1 Baseline Model

IS-ADC builds upon ADC [6], a mask-conditioned latent diffusion framework with teacher-student distillation. ADC is implemented based on the Diffusion Model (LDM) architecture [20]. During training, the teacher receives paired images and masks, whereas the student is conditioned only on masks. The teacher supervises the student via adaptive distillation, and only the student is retained during inference.

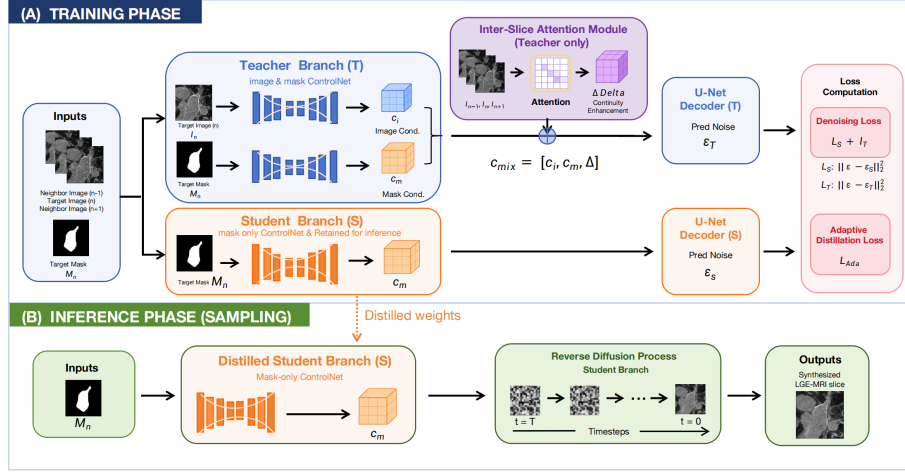


Fig. 1. Architecture of the proposed IS-ADC framework. To improve anatomical consistency in synthetic LGE-MRI generation, neighbouring slices are exploited by a teacher-only inter-slice attention module that captures contextual relationships and generates continuity-enhanced features. These features are integrated with image and mask conditioning in the Adaptively Distilled ControlNet framework during training. At inference, the distilled student branch alone performs mask-conditioned reverse diffusion to synthesise anatomically realistic LGE-MRI slices.

Despite its effectiveness, ADC processes each slice independently and therefore neglects the strong anatomical continuity present in volumetric LGE-MRI. Adjacent slices typically exhibit gradual variations in atrial geometry and tissue boundaries, providing valuable contextual cues that cannot be captured by isolated 2D generation. Consequently, slice-wise synthesis may introduce discontinuities that degrade volumetric coherence. To address this limitation, we propose an Inter-Slice Attention module that employs neighbouring-slice information during teacher-guided training.

2.2 Inter-Slice Attention

To incorporate volumetric contextual information while maintaining the efficiency of 2D latent diffusion, we introduce an Inter-Slice Attention (ISA) module inspired by sparse temporal attention in video diffusion models [13]. Rather than treating each slice independently, ISA enables the target-slice representation to incorporate anatomical context from neighbouring slices.

Given three consecutive slices I_{n-1}, I_n, I_{n+1} , the target slice I_n is processed by the original ADC teacher branch to obtain the image-conditioned feature c_i , which is fused with the mask-conditioned feature c_m as $c_{mix} = c_i + c_m$. In parallel, the neighbouring slices are fed into the proposed ISA module to estimate an inter-slice contextual residual, as shown in Fig. 1 (A) training phase. Let

f_{n-1} , f_n , and f_{n+1} denote the projected latent features of the previous, target, and subsequent slices, respectively. The neighbouring context is constructed by concatenating the adjacent-slice features: $f_N(n) = [f_{n-1}, f_{n+1}]$. The target-slice feature serves as the query (Q), while $f_N(n)$ is used to construct the keys (K) and values (V):

$$Q = W_Q f_n, \quad K = W_K f_N(n), \quad V = W_V f_N(n). \quad (1)$$

where W_Q , W_K , and W_V are learnable projection matrices. The inter-slice contextual feature is computed as

$$A_n = \text{Softmax} \left(\frac{QK^\top}{\sqrt{d}} \right) V, \quad (2)$$

where d is the projected feature dimension. The contextual residual is defined as

$$\Delta_n = A_n - f_n, \quad (3)$$

which represents the contextual feature adjustment derived from neighbouring slices. The residual is then injected into the teacher branch through

$$c'_{mix} = c_{mix} + \alpha \Delta_n, \quad (4)$$

where α is a fixed scaling coefficient controlling the contribution of the inter-slice contextual residual. Based on empirical comparison of $\alpha \in \{0.3, 0.5, 1\}$, we set $\alpha = 1$ in the final model. The refined feature c'_{mix} is subsequently propagated to the teacher decoder for noise prediction and distillation supervision.

No additional supervision is introduced for the ISA module. Instead, it is optimised jointly with the original ADC objective through diffusion denoising and adaptive distillation. Since accurate noise prediction depends on anatomically consistent LGE-MRI representations, the attention mechanism learns to exploit relevant contextual information from neighbouring slices.

During inference, only the distilled student branch is retained for mask-conditioned image synthesis, preserving the computational efficiency of the original ADC framework.

3 Experiments

3.1 Dataset

We evaluated the proposed IS-ADC framework on the STACOM 2018 Left Atrial Segmentation Challenge dataset [17]. The dataset contains 154 late gadolinium-enhanced magnetic resonance imaging (LGE-MRI) volumes with corresponding left atrial (LA) cavity annotations. Following previous studies [16], all scans were acquired with an in-plane resolution of 0.625×0.625 mm and a slice thickness of 1.25 mm. The dataset was randomly divided into 104 training, 20 validation, and 30 testing cases.

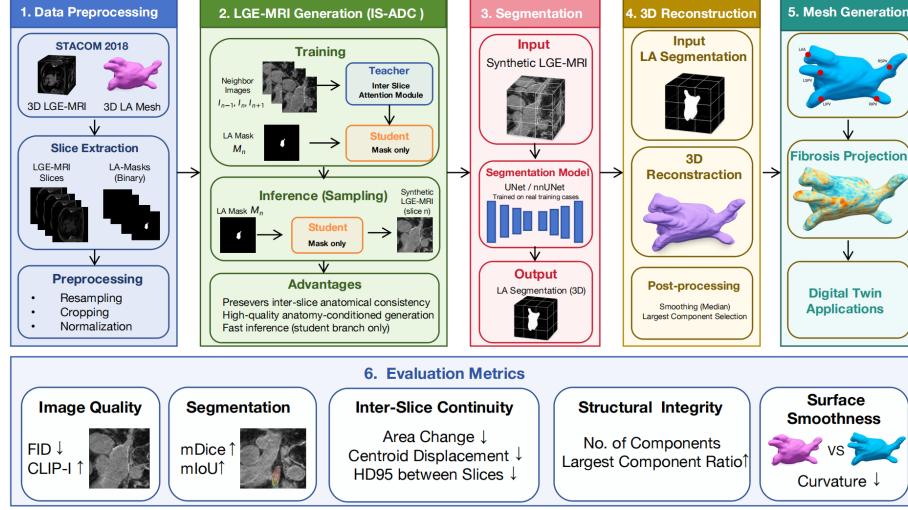


Fig. 2. Overall framework of the proposed pipeline. Following data preprocessing, the proposed IS-ADC model generates synthetic LGE-MRI slices from anatomical masks while incorporating neighbouring-slice information during training to enhance anatomical continuity. The synthesised images are subsequently used for LA segmentation, 3D reconstruction, and mesh generation. Quantitative evaluation is performed using image quality, segmentation, inter-slice continuity, structural integrity, and surface smoothness metrics.

As illustrated in Fig. 2, all volumes underwent a standardised preprocessing pipeline, including slice extraction, mask-guided cropping, resizing, and normalisation. Images were resized to 256×256 pixels and normalised to $[-1, 1]$, while binary mask inputs were normalised to $[0, 1]$.

3.2 Implementation Details

The proposed IS-ADC was trained on a single NVIDIA H200 NVL GPU using the AdamW optimiser with a learning rate of 1×10^{-4} and a batch size of 8 for 14,000 iterations. During training, the teacher branch receives three consecutive slices $\{I_{n-1}, I_n, I_{n+1}\}$ together with the target anatomical mask M_n , whereas the student branch is conditioned only on the target anatomical mask M_n . A prompt dropout rate of 5% was adopted during training. During inference, synthetic images were generated using Denoising Diffusion Implicit Models (DDIM) sampling [21] with 50 denoising steps ($\eta = 0$) and a classifier-free guidance scale [22] of 9.

For comparison, we also implemented a three-slice ADC baseline, denoted as ADC+3-slice, where three consecutive slices $\{I_{n-1}, I_n, I_{n+1}\}$ were concatenated and provided to the ControlNet (teacher branch) as input. Unlike IS-ADC, this baseline does not include the proposed Inter-Slice Attention module. It was de-

Table 1. Comparison of training time.

Model	Time/Iteration (s)↓
ADC	0.779
ADC+3-slice input	0.780
IS-ADC	0.788

Runtime was averaged over 200 training iterations after 30 warm-up iterations under identical experimental settings.

signed to determine whether simply providing neighbouring slices is sufficient to improve inter-slice continuity.

To quantify the additional training overhead introduced by the proposed method, we compared the training time of ADC, ADC+3-slice input, and IS-ADC under identical settings. As shown in Table 1, incorporating three consecutive slices alone incurs only a marginal computational overhead. Compared with ADC, IS-ADC increases the training time per iteration from 0.779 s to 0.788 s, corresponding to an overhead of approximately 1.2%. Since the inter-slice attention module is used only in the teacher branch during training and is discarded during inference, the proposed method introduces no additional inference cost.

3.3 Evaluation Metrics and Methods

Image synthesis quality was evaluated using Fréchet Inception Distance (FID) [18] and CLIP Image Similarity (CLIP-I) [19].

To assess segmentation-based anatomical fidelity, U-Net [24] and nnU-Net [23] trained on the real training set were evaluated on synthetic test images. Segmentation performance was quantified using the Dice score (Dice) and Intersection-over-Union (IoU).

To investigate whether improved inter-slice consistency translated into superior volumetric coherence, predicted segmentation slices were stacked into three-dimensional left atrial volumes using the reconstruction pipeline shown in Fig. 2.

As illustrated in Fig. 2, fibrosis distributions were projected onto reconstructed meshes following previous work [4, 25, 26] as a representative downstream application. Binary masks were median-filtered before reconstruction and meshes were Laplacian smoothed. Quantitative evaluation of fibrosis realism, however, is beyond the scope of the present work and will be investigated in future studies.

Volumetric anatomical fidelity was evaluated from four aspects: area consistency, spatial alignment, topological integrity, and surface smoothness. Area consistency was quantified using the relative area variation between adjacent slices (ΔA). Spatial alignment was assessed using centroid displacement (CD), while boundary continuity was evaluated using the inter-slice 95th percentile Hausdorff distance (HD95). Topological integrity was characterised by the number of connected components (CC) and the largest connected-component ratio (LCR). Surface smoothness was quantified using the mean absolute curvature (MAC).

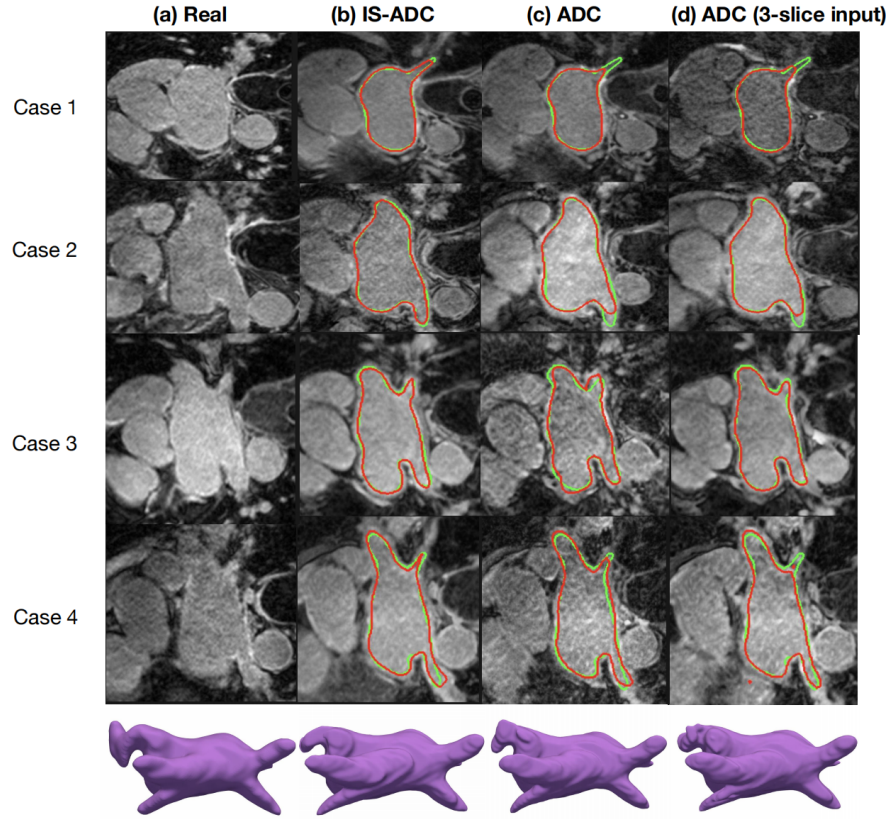


Fig. 3. Qualitative comparison of real and synthetic LGE-MRI images and their reconstructed LA meshes. Green and red contours indicate the reference and predicted segmentations, respectively.

3.4 Results

Figure 3 presents representative comparisons of real and synthetic LGE-MRI images. Green and red contours denote the reference and predicted segmentation boundaries, respectively. Compared with ADC and ADC+3-slice, IS-ADC generally exhibits improved boundary agreement, including in geometrically complex regions around the pulmonary veins and atrial appendage. These improvements are further reflected in the reconstructed meshes, where IS-ADC produces more anatomically plausible left atrial geometries with fewer local surface irregularities.

Table 2 summarises image synthesis quality and downstream segmentation performance. IS-ADC achieved the lowest FID (37.06) and highest CLIP-I (0.94), indicating improved image realism and semantic consistency. It also obtained the best Dice and mIoU scores for both U-Net and nnU-Net. A two-sided paired

Table 2. Image synthesis quality and segmentation-based anatomical fidelity.

Model	FID↓	CLIP-I↑	U-Net (%)		nnU-Net (%)	
			Dice↑	mIoU↑	Dice↑	mIoU↑
Real	—	—	90.79	83.29	93.05	87.09
ADC	45.95	0.93	91.34	84.11	93.38	87.64
ADC+3-slice input	49.83	0.93	91.03	83.59	93.51	87.87
IS-ADC	37.06	0.94	91.76*	84.83*	94.43*	89.49*

* Significant improvement over ADC and ADC+3-slice input ($p < 0.01$, two-sided Wilcoxon signed-rank test).

Table 3. Quantitative evaluation of volumetric anatomical consistency and reconstruction quality.

Model	ΔA ↓	CD (mm) ↓	HD95 (mm) ↓	CC ↓	LCR ↑	MAC ↓
Real	0.095	1.585	4.159	1.000	1.000	0.081
ADC	0.214	3.799	9.433	1.200	0.998	0.093
ADC+3-slice input	0.228	3.656	8.718	1.070	0.999	0.095
IS-ADC	0.188*	2.793*	6.709*	1.033	1.000	0.087*

* Significant improvement over both ADC and ADC+3-slice input (two-sided paired Wilcoxon signed-rank test, $p < 0.01$).

Wilcoxon signed-rank test on case-level segmentation scores showed that these segmentation improvements were significant compared with both ADC and ADC+3-slice input ($p < 0.01$). In contrast, simply adding neighbouring slices did not improve synthesis quality.

Table 3 further evaluates volumetric anatomical consistency. Among the synthesis methods, IS-ADC achieved the lowest ΔA , CD, and HD95, reducing them by 12.1%, 26.5%, and 28.9% compared with ADC. The improvements in ΔA , CD, HD95, and MAC were also significant under a two-sided paired Wilcoxon signed-rank test on case-level measurements ($p < 0.01$). IS-ADC additionally produced fewer disconnected components and achieved an LCR of 1.000, indicating improved topological integrity and reconstruction quality.

To assess the utility of synthetic images for segmentation model training, nnU-Net was trained under five data-composition settings: (1) 100% Synthetic, (2) 100% Real, (3) 50% Real + 50% Synthetic, (4) 100% Real + 50% Synthetic, and (5) 100% Real + 100% Synthetic, where each synthetic image was generated from the anatomical mask corresponding to a real training image. As shown in Fig. 4, incorporating synthetic images slightly improved segmentation performance over training with real images alone. The best performance was achieved using setting (5), reaching 93.11% Dice and 87.19% mIoU, compared with 93.05% Dice and 87.09% mIoU using setting (2). Furthermore, training solely on synthetic images still achieved 92.92% Dice and 86.86% mIoU, demonstrating that the generated images preserve sufficient anatomical fidelity for effective segmentation. These results indicate that IS-ADC-generated images can serve as useful complementary training data for downstream medical image analysis.

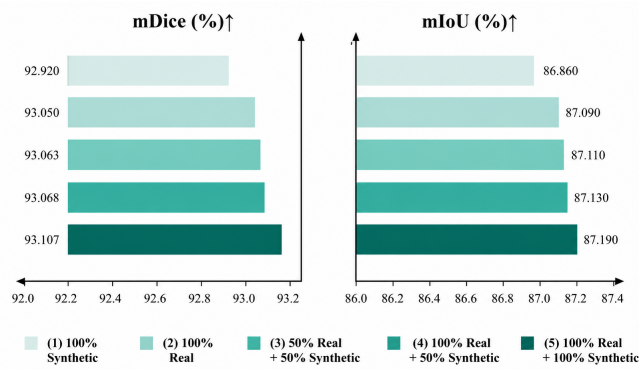


Fig. 4. Quantitative comparison on data augmentation.

4 Conclusion

This paper presented IS-ADC, an inter-slice-attention-enhanced diffusion framework for left atrial LGE-MRI synthesis. By modelling dependencies between neighbouring slices during training, IS-ADC improves anatomical consistency while retaining the inference efficiency of a 2D diffusion model. Compared with the original ADC and ADC+3-slice, IS-ADC achieved better image fidelity, segmentation performance, inter-slice continuity, and 3D reconstruction quality.

A limitation of this study is that the proposed framework was evaluated on a single left atrial LGE-MRI dataset. Future work will investigate generalisation across datasets, imaging protocols, and modalities and explore the integration of generated anatomies into patient-specific cardiac digital twin pipelines for electrophysiological modelling and virtual clinical trials. In addition, this work focuses primarily on anatomical continuity and geometric consistency. Evaluating the pathological realism of synthesised fibrosis patterns and their influence on downstream electrophysiological simulations remains an important direction for future research.

5 Acknowledgements

This work utilised Queen Mary’s Apocrita HPC facility, supported by QMUL Research-IT (<http://doi.org/10.5281/zenodo.438045>). This research was funded by a UKRI Future Leaders Fellowship (MR/W004720/1). We also acknowledge funding from the National Institute for Health Research Barts Biomedical Research Centre (NIHR203330). This study was also funded by the British Heart Foundation Clinical Intermediate Fellowship (FS/ICRF/22/26034).

6 Disclosure of Interests

The authors declare no competing interests.

References

1. Nattel, S., Burstein, B., Dobrev, D.: Atrial remodelling and atrial fibrillation: mechanisms and implications. *Circ. Arrhythm. Electrophysiol.* **1**(1), 62–73 (2008)
2. Hersi, A., Mitchell, L.B., Wyse, D.G.: Atrial fibrillation: Challenges and opportunities. *Can. J. Cardiol.* **22**(Suppl. C), 21C (2006)
3. Wink, K.C.J., Roelofs, E., Simone II, C.B., Dechambre, D., Santiago, A., Van Der Stoep, J., et al.: Photons, protons or carbon ions for stage I non-small cell lung cancer: Results of the multicentric ROCOCO in silico study. *Radiother. Oncol.* **128**(1), 139–146 (2018)
4. Solís-Lemus, J.A., Baptiste, T., Barrows, R., Sillett, C., Gharaviri, A., Raffaele, G., Razeghi, O., Strocchi, M., Sim, I., Kotadia, I., et al.: Evaluation of an open-source pipeline to create patient-specific left atrial models: A reproducibility study. *Comput. Biol. Med.* **162**, 107009 (2023)
5. Zhang, L., Rao, A., Agrawala, M.: Adding conditional control to text-to-image diffusion models. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 3836–3847 (2023)
6. Qiu, K., Zhou, Z., Guo, Y.: Adaptively distilled ControlNet: Accelerated training and superior sampling for medical image synthesis. In: *International Conference on Medical Image Computing and Computer-Assisted Intervention (MICCAI)*, LNCS, pp. 55–65. Springer, Cham (2025)
7. Müller-Franzes, G., Niehues, J.M., Khader, F., Arasteh, S.T., Haarbuerger, C., Kuhl, C., Wang, T., Han, T., Nolte, T., Nebelung, S., et al.: A multimodal comparison of latent denoising diffusion probabilistic models and generative adversarial networks for medical image synthesis. *Sci. Rep.* **13**(1), 12098 (2023)
8. Khader, F., Müller-Franzes, G., Tayebi Arasteh, S., Han, T., Haarbuerger, C., Schulze-Hagen, M., Schad, P., Engelhardt, S., Baekler, B., Foersch, S., et al.: Denoising diffusion probabilistic models for 3D medical image generation. *Sci. Rep.* **13**(1), 7303 (2023)
9. Kumar, A., Jiang, H., Imran, M., Valdes, C., Leon, G., Kang, D., Nataraj, P., Zhou, Y., Weiss, M.D., Shao, W.: A flexible 2.5D medical image segmentation approach with in-slice and cross-slice attention. *Comput. Biol. Med.* **182**, 109173 (2024)
10. Zhang, J., Xie, Y., Wang, Y., Xia, Y.: Inter-slice context residual learning for 3D medical image segmentation. *IEEE Trans. Med. Imaging* **40**(2), 661–672 (2020)
11. Zhu, L., Xue, Z., Jin, Z., Liu, X., He, J., Liu, Z., Yu, L.: Make-a-volume: Leveraging latent diffusion models for cross-modality 3D brain MRI synthesis. In: *Medical Image Computing and Computer Assisted Intervention – MICCAI 2023*, LNCS, pp. 592–601. Springer, Cham (2023)
12. Hu, Y., Kothapalli, S.V.V.N., Gan, W., Sukstanskii, A.L., Wu, G.F., Goyal, M., Yablonskiy, D.A., Kamilov, U.S.: DiffGEPCI: 3D MRI synthesis from mGRE signals using 2.5D diffusion model. In: *2024 IEEE International Symposium on Biomedical Imaging (ISBI)*, pp. 1–4. IEEE (2024)
13. Wu, J.Z., Ge, Y., Wang, X., Lei, S.W., Gu, Y., Shi, Y., Hsu, W., Shan, Y., Qie, X., Shou, M.Z.: Tune-a-video: One-shot tuning of image diffusion models for text-to-video generation. In: *Proceedings of the IEEE/CVF International Conference on Computer Vision (ICCV)*, pp. 7623–7633 (2023)
14. Beetz, M., Banerjee, A., Grau, V.: Reconstructing 3D cardiac anatomies from misaligned multi-view magnetic resonance images with mesh deformation U-Nets. In: *Geometric Deep Learning in Medical Image Analysis*, pp. 3–14 (2022)

15. Roney, C.H., Beach, M.L., Mehta, A.M., Sim, I., Corrado, C., Bendikas, R., Solis-Lemus, J.A., Razeghi, O., Whitaker, J., O'Neill, L., et al.: In silico comparison of left atrial ablation techniques that target the anatomical, structural, and electrical substrates of atrial fibrillation. *Front. Physiol.* **11**, 1145 (2020)
16. McGann, C., Akoum, N., Patel, A., Kholmovski, E., Revelo, P., Damal, K., Wilson, B., Cates, J., Harrison, A., Ranjan, R., et al.: Atrial fibrillation ablation outcome is predicted by left atrial remodeling on MRI. *Circ. Arrhythm. Electrophysiol.* **7**(1), 23–30 (2014)
17. Xiong, Z., Xia, Q., Hu, Z., Huang, N., Bian, C., Zheng, Y., Vesal, S., Ravikumar, N., Maier, A., Yang, X., et al.: A global benchmark of algorithms for segmenting the left atrium from late gadolinium-enhanced cardiac magnetic resonance imaging. *Med. Image Anal.* **67**, 101832 (2021)
18. Heusel, M., Ramsauer, H., Unterthiner, T., Nessler, B., Hochreiter, S.: GANs trained by a two time-scale update rule converge to a local Nash equilibrium. In: *Advances in Neural Information Processing Systems (NeurIPS)* (2017)
19. Ruiz, N., Li, Y., Jampani, V., Pritch, Y., Rubinstein, M., Aberman, K.: Dream-Booth: Fine tuning text-to-image diffusion models for subject-driven generation. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 22500–22510 (2023)
20. Rombach, R., Blattmann, A., Lorenz, D., Esser, P., Ommer, B.: High-resolution image synthesis with latent diffusion models. In: *Proceedings of the IEEE/CVF Conference on Computer Vision and Pattern Recognition (CVPR)*, pp. 10684–10695 (2022)
21. Song, J., Meng, C., Ermon, S.: Denoising diffusion implicit models. In: *International Conference on Learning Representations (ICLR)* (2021)
22. Ho, J., Salimans, T.: Classifier-free diffusion guidance. In: *NeurIPS 2021 Workshop on Deep Generative Models and Downstream Applications* (2021)
23. Isensee, F., Jaeger, P.F., Kohl, S.A.A., Petersen, J., Maier-Hein, K.H.: nnU-Net: A self-configuring method for deep learning-based biomedical image segmentation. *Nat. Methods* **18**(2), 203–211 (2021)
24. Ronneberger, O., Fischer, P., Brox, T.: U-Net: Convolutional networks for biomedical image segmentation. In: *Medical Image Computing and Computer-Assisted Intervention – MICCAI 2015, LNCS*, vol. 9351, pp. 234–241. Springer, Cham (2015)
25. Zolotarev, A.M., Johnson, K., Mohammad, Y., Alwazzan, O., Slabaugh, G., Roney, C.H.: Synthetic fibrosis distributions for data augmentation in predicting atrial fibrillation ablation outcomes: An in silico study. *Front. Cardiovasc. Med.* **12**, 1512356 (2025)
26. Roney, C.H., Sim, I., Yu, J., Beach, M., Mehta, A., Alonso Solis-Lemus, J., Kotadia, I., Whitaker, J., Corrado, C., Razeghi, O., et al.: Predicting atrial fibrillation recurrence by combining population data and virtual cohorts of patient-specific left atrial models. *Circ. Arrhythm. Electrophysiol.* **15**(2), e010253 (2022)